

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2003-263279

(43)Date of publication of application : 19.09.2003

(51)Int.Cl.

G06F 3/06
G06F 12/08

(21)Application number : 2003-086908

(71)Applicant : HITACHI LTD

(22)Date of filing : 18.09.1998

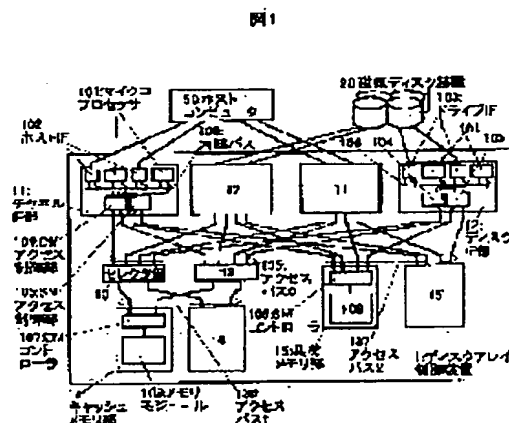
(72)Inventor : FUJIMOTO KAZUHISA
TANAKA ATSUSHI
FUJIBAYASHI AKIRA
KANAI HIROKI
MINOWA NOBUYUKI

(54) DISK ARRAY CONTROL APPARATUS

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a disk array control apparatus which realizes a high throughput and short response time with consideration of properties of data to be stored into cache memory and shared memory, and access properties for these memories.

SOLUTION: The disk array control apparatus has a plurality of channels IF, a plurality of disks IF, a cache memory and a shared memory. The connection form between the plurality of channels IF and the plurality of disks IF, and the cache memory is different from that between the plurality of channels IF and the shared memory, and the plurality of disks IF and the shared memory. In this disk array control apparatus, access paths to the cache memory and the shared memory are increased to improve throughput, and on the other hand, access time to the shared memory can be shortened.



LEGAL STATUS

[Date of request for examination]

25.08.2005

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C): 1998,2003 Japan Patent Office

THIS PAGE BLANK (USPTO)

(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号
特開2003-263279
(P2003-263279A)

(43)公開日 平成15年 9 月19日 (2003. 9. 19)

(51)Int.Cl. ⁷	識別記号	F I	テ-マコ-ト*(参考)
G 0 6 F 3/06	3 0 2	C 0 6 F 3/06	3 0 2 A 5 B 0 0 5
	5 4 0		5 4 0 5 B 0 6 5
12/08	5 1 1	12/08	5 1 1 Z
	5 4 1		5 4 1 Z
	5 5 7		5 5 7
審査請求 未請求 請求項の数26 O L (全 20 頁)			

(21)出願番号	特願2003-86908(P2003-86908)	(71)出願人	000005108
(62)分割の表示	特願平10-264286の分割		株式会社日立製作所
(22)出願日	平成10年 9 月18日 (1998. 9. 18)		東京都千代田区神田駿河台四丁目 6 番地
		(72)発明者	藤本 和久
			東京都国分寺市東恋ヶ窪一丁目280番地
			株式会社日立製作所中央研究所内
		(72)発明者	田中 淳
			東京都国分寺市東恋ヶ窪一丁目280番地
			株式会社日立製作所中央研究所内
		(74)代理人	100075096
			弁理士 作田 康夫

最終頁に続く

(54)【発明の名称】 ディスクアレイ制御装置

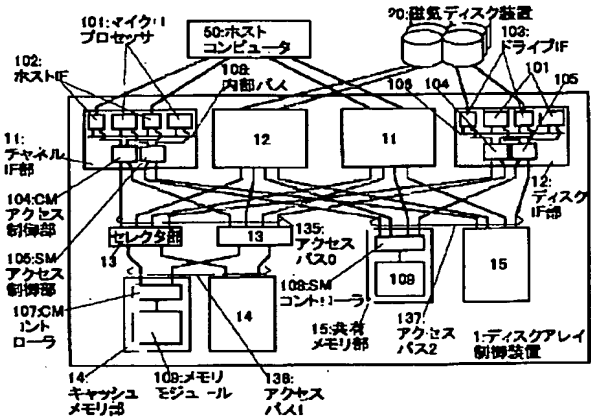
(57)【要約】

【課題】 キャッシュメモリ及び共有メモリに格納されるデータの特性及びこれらのメモリへのアクセス特性を考慮した、スループットが高く、かつ、応答時間の短いディスクアレイ制御装置を提供することにある。

【解決手段】 上記課題は、複数のチャンネル I F 部と、複数のディスク I F 部と、キャッシュメモリ部と、共有メモリ部とを有し、前記複数のチャンネル I F 部及び前記複数のディスク I F 部と前記キャッシュメモリ部との間の接続形式が、前記複数のチャンネル I F 部及び前記複数のディスク I F 部と前記共有メモリ部との間の接続形式と異なることを特徴とするディスクアレイ制御装置により達成される。

【効果】 キャッシュメモリ及び共有メモリへのアクセスパスを増やしスループットを高くする一方、共有メモリへのアクセス時間を短くすることができる。

図 1



(2) 003-263279 (P2003-263279A)

【特許請求の範囲】

【請求項1】ホストコンピュータとのインターフェースを有する複数のチャネルインターフェース部と、磁気ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、前記磁気ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリ部と、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部との間のデータ転送に関する制御情報を格納する共有メモリ部とを有し、各チャネルインターフェース部は、前記ホストコンピュータとのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行し、各ディスクインターフェース部は、前記磁気ディスク装置とのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行するディスクアレイ制御装置において、前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記キャッシュメモリ部との間の接続形式が、前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記共有メモリ部との間の接続形式と異なることを特徴とするディスクアレイ制御装置。

【請求項2】前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記キャッシュメモリ部との間はセレクト部を介して接続され、前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記共有メモリ部との間は、それぞれセレクト部を介さずに接続されていることを特徴とする請求項1に記載のディスクアレイ制御装置。

【請求項3】前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記キャッシュメモリ部との間はセレクト部を介して接続され、前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記共有メモリ部との間はそれぞれは1対1接続されていることを特徴とする請求項1に記載のディスクアレイ制御装置。

【請求項4】前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記キャッシュメモリ部との間はセレクト部を介して接続され、前記複数のチャネルインターフェース部、前記複数のディスクインターフェース部、及び前記共有メモリ部は共有バスに接続されていることを特徴とする請求項1に記載のディスクアレイ制御装置。

【請求項5】前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記キャッシュメモリ部との間はスイッチを用いた相互結合網によって接続され、前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記共有メモリ部との間はそれぞれ直接接続されていることを特徴とする請求項1に記載のディスクアレイ制御装置。

【請求項6】前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記キャッシュメモリ部との間はスイッチを用いた相互結合網によって接続され、前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記共有メモリ部との間はそれぞれ1対1接続されていることを特徴とする請求項1に記載のディスクアレイ制御装置。

【請求項7】ホストコンピュータとのインターフェースを有する複数のチャネルインターフェース部と、磁気ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、前記複数のチャネルインターフェース部と前記複数のディスクインターフェース部とに接続され、前記磁気ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリ部と、前記複数のチャネルインターフェース部と前記複数のディスクインターフェース部とに接続され、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部との間のデータ転送に関する制御情報を格納する前記共有メモリ部とを有する制御装置において、前記複数のチャネルインターフェース部と前記複数のディスクインターフェース部から前記キャッシュメモリ部に接続されるアクセスパスの本数は、前記複数のチャネルインターフェース部と前記複数のディスクインターフェース部から前記共有メモリ部に接続されるアクセスパスの本数より少ないことを特徴とするディスクアレイ制御装置。

【請求項8】セレクト部をさらに有し、前記各チャネルインターフェース部及び前記各ディスクインターフェース部と前記セレクト部とは、それぞれアクセスパスにより1対1に接続され、前記セレクト部と前記キャッシュメモリ部とは、アクセスパスにより接続され、前記セレクト部を介して前記キャッシュメモリ部と接続された前記チャネルインターフェース部及び前記ディスクインターフェース部と前記セレクト部とを接続する前記アクセスパスの本数は、前記セレクト部と前記メモリ部を接続する前記アクセスパスの本数より多く、前記各チャネルインターフェース部及び前記各ディスクインターフェース部と前記共有メモリ部との間はそれぞれアクセスパスにより1対1に接続されていることを特徴とする請求項7に記載のディスクアレイ制御装置。

【請求項9】ホストコンピュータとの複数のインターフェース部と、磁気ディスク装置との複数のインターフェース部と、前記磁気ディスク装置のデータを一時的に格納する物理的に独立したキャッシュメモリ部と、制御情報を格納する物理的に独立した共有メモリ部とを有し、前記ホストコンピュータとの各インターフェース部及び前記磁気ディスク装置との各インターフェース部は、それぞれマイクロプロセッサと、前記キャッシュメモリ部へのアクセスを制御するキャッシュメモリアクセス制御部と、前記共有メモリ部へのアクセスを制御する共有メ

!(3) 003-263279 (P2003-263279A)

モリアクセス制御部とを有するディスクアレイ制御装置において、前記各キャッシュメモリアクセス制御部と前記キャッシュメモリ部との間は、セクタ部を介してアクセスパスにより接続されており、前記各共有メモリアクセス制御部と前記共有メモリ部との間は、それぞれアクセスパスにより1対1接続されていることを特徴とするディスクアレイ制御装置。

【請求項10】ホストコンピュータとの複数のインターフェース部と、磁気ディスク装置との複数のインターフェース部と、前記磁気ディスク装置のデータを一時的に格納する物理的に独立したキャッシュメモリ部と、制御情報を格納する物理的に独立した共有メモリ部とを有し、前記ホストコンピュータとの各インターフェース部及び前記磁気ディスク装置との各インターフェース部は、それぞれマイクロプロセッサと、前記キャッシュメモリ部へのアクセスを制御するキャッシュメモリアクセス制御部と、前記共有メモリ部へのアクセスを制御する共有メモリアクセス制御部とを有するディスクアレイ制御装置において、前記各キャッシュメモリアクセス制御部と前記キャッシュメモリ部との間は、スイッチを用いた相互結合網によって接続されており、前記各共有メモリアクセス制御部と前記共有メモリ部との間は、それぞれアクセスパスにより1対1接続されていることを特徴とするディスクアレイ制御装置。

【請求項11】前記各キャッシュメモリアクセス制御部と前記キャッシュメモリ部との間の前記アクセスパスの帯域幅を、前記ホストコンピュータと前記ホストコンピュータとのインターフェース部との間の最大の全帯域幅の2倍以上としたことを特徴とする請求項9または請求項10の何れかに記載のディスクアレイ制御装置。

【請求項12】前記共有メモリ部及び前記キャッシュメモリ部を複数有し、該複数の共有メモリ部及び前記複数のキャッシュメモリ部はそれぞれ二重化されていることを特徴とする請求項1乃至請求項11の何れかに記載のディスクアレイ制御装置。

【請求項13】ホストコンピュータとの複数のチャンネルインターフェースと、ディスク装置との複数のディスクインターフェースと、前記チャンネルインターフェースと前記ディスクインターフェースを1対1接続可能な第一のアクセスパスと、前記チャンネルインターフェースと前記ディスクインターフェースを接続する第二のアクセスパスを備え、前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータは前記第一のアクセスパスを介して伝送され、また、データ伝送に関する制御情報は前記第二のアクセスパスを介して伝送され、さらに、前記第一のアクセスパスの帯域幅は、前記第二のアクセスパスの帯域幅よりも広いことを特徴とするディスクアレイ制御装置。

【請求項14】ホストコンピュータとの複数のチャンネルインターフェースと、ディスク装置との複数のディスク

インターフェースと、前記チャンネルインターフェースと前記ディスクインターフェースを1対1接続可能なデータ用ネットワークと、前記チャンネルインターフェースと前記ディスクインターフェースを接続する制御情報用ネットワークを備え、前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータは前記データ用ネットワークを介して伝送され、データ伝送に関する制御情報は前記制御情報用ネットワークを介して、前記データ用ネットワークとは独立に、伝送されることを特徴とするディスクアレイ制御装置。

【請求項15】ホストコンピュータとの複数のチャンネルインターフェースと、ディスク装置との複数のディスクインターフェースを備え、前記チャンネルインターフェースは第一及び第二のアクセス制御部を備え、前記ディスクインターフェースは第三及び第四のアクセス制御部を備え、前記第一のアクセス制御部と前記第三のアクセス制御部は1対1接続可能であり、前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータは、前記第一のアクセス制御部及び前記第三のアクセス制御部を介して伝送され、データ伝送に関する制御情報は、前記第二のアクセス制御部及び前記第四のアクセス制御部を介して伝送されることを特徴とするディスクアレイ制御装置。

【請求項16】ホストコンピュータとの複数のチャンネルインターフェースと、ディスク装置との複数のディスクインターフェースと、前記複数のチャンネルインターフェースと前記複数のディスクインターフェースを接続するスイッチを備え、データ伝送に関する制御情報は、前記スイッチを介して伝送されることを特徴とするディスクアレイ制御装置。

【請求項17】ホストコンピュータとの複数のチャンネルインターフェースと、ディスク装置との複数のディスクインターフェースと、前記複数のチャンネルインターフェースと前記複数のディスクインターフェースを接続する第一のスイッチと、前記複数のチャンネルインターフェースと前記複数のディスクインターフェースを接続する第二のスイッチを備え、前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータは、前記第一のスイッチを介して伝送され、データ伝送に関する制御情報は、前記第二のスイッチを介して伝送されることを特徴とするディスクアレイ制御装置。

【請求項18】ホストコンピュータとの複数のチャンネルインターフェースと、ディスク装置との複数のディスクインターフェースと、前記チャンネルインターフェースと第一のアクセスパスを介して接続され、前記ディスクインターフェースと第二のアクセスパスを介して接続されたスイッチを備え、前記チャンネルインターフェースは、データ伝送に関する制御情報を前記第一のパスを介して前記複数のディスクインターフェースと通信し、前記ディスクインターフェースは、データ伝送に関する制御情

(4) 003-263279 (P2003-263279A)

報を前記第二のパスを介して前記複数のチャネルインターフェースと通信することを特徴とするディスクアレイ制御装置。

【請求項19】ディスク装置に接続されたディスクアレイ制御装置であって、外部装置との第一のインターフェースと、外部装置との第二のインターフェースと、前記第一のインターフェースと前記第二のインターフェースを1対1接続可能な第一のアクセスパスと、前記第一のインターフェースと前記第二のインターフェースを接続する第二のアクセスパスを備え、前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータは、前記第一のアクセスパスを介して伝送され、データ伝送に関する制御情報は、前記第二のアクセスパスを介して伝送され、前記第一のアクセスパスの帯域幅は前記第二のアクセスパスの帯域幅よりも広いことを特徴とするディスクアレイ制御装置。

【請求項20】ディスク装置に接続されたディスクアレイ制御装置であって、外部装置との第一のインターフェースと、外部装置との第二のインターフェースと、前記第一のインターフェースと前記第二のインターフェースを1対1接続可能なデータ用ネットワークと、前記第一のインターフェースと前記第二のインターフェースを接続する制御情報用ネットワークを備え、前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータは、前記データ用ネットワークを介して伝送され、データ伝送に関する制御情報は前記制御情報用ネットワークを介して、前記データ用ネットワークとは独立に、伝送されることを特徴とするディスクアレイ制御装置。

【請求項21】ディスク装置に接続されたディスクアレイ制御装置であって、外部装置との第一のインターフェースと、外部装置との第二のインターフェースと、前記第一のインターフェースと前記第二のインターフェースを接続する第一のスイッチと、前記第一のインターフェースと前記第二のインターフェースを接続する第二のスイッチと、前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータは前記第一のスイッチを介して伝送され、データ伝送に関する制御情報は前記第二のスイッチを介して伝送されることを特徴とするディスクアレイ制御装置。

【請求項22】ディスク装置に接続されたディスクアレイ制御装置であって、外部装置との複数の第一のインターフェースと、外部装置との複数の第二のインターフェースと、前記複数の第一のインターフェースと第一のアクセスパスを介して接続され、前記複数の第二のインターフェースと第二のアクセスパスを介して接続されたスイッチを備え、前記第一のインターフェースは、データ伝送に関する制御情報を前記第一のアクセスパスを介して、前記複数の第二のインターフェースと通信し、前記第二のインターフェースは、データ伝送に関する制御情

報を前記第二のアクセスパスを介して、前記複数の第一のインターフェースと通信することを特徴とするディスクアレイ制御装置。

【請求項23】ホストコンピュータとのインターフェースを有する複数のチャネルインターフェース部と、ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、前記チャネルインターフェース部と前記ディスクインターフェース部を1対1接続可能な第一のアクセスパスと、前記チャネルインターフェース部と前記ディスクインターフェース部を接続する第二のアクセスパスを備え、前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータは、前記第一のアクセスパスを介して伝送され、データ伝送に関する制御情報は、前記第二のアクセスパスを介して伝送され、前記第一のアクセスパスの帯域幅は前記第二のアクセスパスの帯域幅よりも広いことを特徴とするディスクアレイ制御装置。

【請求項24】ホストコンピュータとのインターフェースを有する複数のチャネルインターフェース部と、ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、前記チャネルインターフェース部と前記ディスクインターフェース部を1対1接続可能なデータ用ネットワークと、前記チャネルインターフェース部と前記ディスクインターフェース部を接続する制御情報用ネットワークを備え、前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータは、前記データ用ネットワークを介して伝送され、データ伝送に関する制御情報は前記制御情報用ネットワークを介して、前記データ用ネットワークとは独立に、伝送されることを特徴とするディスクアレイ制御装置。

【請求項25】ホストコンピュータとのインターフェースを有する複数のチャネルインターフェース部と、ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、前記チャネルインターフェース部と前記ディスクインターフェース部を接続する第一のスイッチと、前記チャネルインターフェース部と前記ディスクインターフェース部を接続する第二のスイッチと、前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータは前記第一のスイッチを介して伝送され、データ伝送に関する制御情報は前記第二のスイッチを介して伝送されることを特徴とするディスクアレイ制御装置。

【請求項26】ホストコンピュータとのインターフェースを有する複数のチャネルインターフェース部と、ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、前記複数のチャネルインターフェース部と第一のアクセスパスを介して接続され、前記複数のディスクインターフェース部と第二のアクセスパスを介して接続されたスイッチを備え、前記チャネルイ

(5) 003-263279 (P2003-263279A)

ンターフェース部は、データ伝送に関する制御情報を前記第一のアクセスパスを介して、前記複数のディスクインターフェース部と通信し、前記ディスクインターフェース部は、データ伝送に関する制御情報を前記第二のアクセスパスを介して、前記複数のチャンネルインターフェース部と通信することを特徴とするディスクアレイ制御装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、データを複数の磁気ディスク装置に格納するディスクアレイ装置の制御装置に関する。

【0002】

【従来の技術】半導体記憶装置を記憶媒体とするコンピュータの主記憶のI/O性能に比べて、磁気ディスクを記憶媒体とするディスクサブシステム（以下「サブシステム」という。）のI/O性能は3～4桁程度小さく、従来からこの差を縮めること、すなわちサブシステムのI/O性能を向上させる努力がなされている。サブシステムのI/O性能を向上させるための1つの方法として、複数の磁気ディスク装置でサブシステムを構成し、データを複数の磁気ディスク装置に格納する、いわゆるディスクアレイと呼ばれるシステムが知られている。

【0003】図2は、従来のディスクアレイの構成を示す。ホストコンピュータ50とディスクアレイ制御装置2との間のデータ転送を実行する複数のチャンネルIF部11と、磁気ディスク装置20とディスクアレイ制御装置2間のデータ転送を実行する複数のディスクIF部12と、磁気ディスク装置20のデータを一時的に格納するキャッシュメモリ部14と、ディスクアレイ制御装置2に関する制御情報（例えば、チャンネルIF部及びディスクIF部とキャッシュメモリ部14との間のデータ転送制御に関する情報）を格納する共有メモリ部15とを備え、キャッシュメモリ部14および共有メモリ部15は全てのチャンネルIF部11及びディスクIF部12からアクセス可能な構成となっている。このディスクアレイでは、チャンネルIF部11及びディスクIF部12と共有メモリ部15との間、及び、チャンネルIF部11及びディスクIF部12とキャッシュメモリ部14との間は1対1に接続される。以下、このような接続形態をスター接続と呼ぶ。

【0004】チャンネルIF部11は、ホストコンピュータ50と接続するためのインターフェース及びホストコンピュータ50に対する入出力を制御するマイクロプロセッサ（図示せず）を有している。また、ディスクIF部12は、磁気ディスク装置20と接続するためのインターフェース及び磁気ディスク装置20に対する入出力を制御するマイクロプロセッサ（図示せず）を有している。また、ディスクIF部12は、RAID機能の実行も行う。

【0005】図3は、他の従来のディスクアレイの構成を示す。ホストコンピュータ50とディスクアレイ制御装置3間のデータ転送を実行する複数のチャンネルIF部11と、磁気ディスク装置20とディスクアレイ制御装置3間のデータ転送を実行する複数のディスクIF部12と、磁気ディスク装置20のデータを一時的に格納するキャッシュメモリ部14と、ディスクアレイ制御装置3に関する制御情報（例えば、チャンネルIF部及びディスクIF部とキャッシュメモリ部14との間のデータ転送制御に関する情報）を格納する共有メモリ部15を備え、各チャンネルIF部11及びディスクIF部12と共有メモリ部15間には共有バス130で接続され、各チャンネルIF部11及びディスクIF部12とキャッシュメモリ部14間には共有バス131で接続される。以下、このような接続形態を共有バス接続形式と呼ぶ。

【0006】

【発明が解決しようとする課題】ディスクアレイのアーキテクチャーをスケーラブルなものとするには、ディスク制御装置に接続するディスク容量（論理ボリューム数）に応じ、ディスクIF部を増設し、また、必要なホストコンピュータとのチャンネル数に応じて、ディスクアレイ制御装置内のチャンネルIF部を増設する必要がある。しかし、第3図に示した共有バス接続形式のディスクアレイ制御装置では、一旦実装した共有バスの転送能力をチャンネルIF部、ディスクIF部の増設に応じて変更することはできないので、チャンネルIF部、ディスクIF部の増設に柔軟に対応することが困難である。

【0007】また、第3図に示した共有バス接続形式のディスクアレイ制御装置では、ホストコンピュータとディスクアレイ制御装置との間のデータ転送を実行するチャンネルIF部に設けられたホストコンピュータに対する入出力を制御するマイクロプロセッサ、及び磁気ディスク装置とディスクアレイ制御装置との間のデータ転送を実行するディスクIF部に設けられた磁気ディスク装置に対する入出力を制御するマイクロプロセッサに高性能なプロセッサを使用した場合に、これらのプロセッサの性能に比べて、共有バスの転送能力がボトルネックになり、プロセッサの高速化に追従することが困難となる。

【0008】さらに、第3図に示した共有バス接続形式では、共有バスに接続された複数のチャンネルIF部（または複数のディスクIF部）の何れかのチャンネルIF部（またはディスクIF部）に障害が発生した場合に、障害の発生したチャンネルIF部（またはディスクIF部）を特定することが困難である。

【0009】一方、第2図に示したスター接続形式のディスクアレイ制御装置では、共有メモリ部またはキャッシュメモリ部に接続したアクセスパス数に比例して内部バス性能が増加させることができるので、チャンネルIF部、ディスクIF部の増設、または使用するプロセッサの性能に応じて、内部バス性能を増加させることが可能

(6) 003-263279 (P2003-263279A)

である。また、チャンネルＩＦ部及びディスクＩＦ部とキャッシュメモリ部との間、チャンネルＩＦ部及びディスクＩＦ部と共有メモリ部との間がスター接続されているため、障害の発生したチャンネルＩＦ部（またはディスクＩＦ部）を特定することも容易である。

【0010】スター接続形式のディスクアレイ制御装置では、搭載されるチャンネルＩＦ部またはディスクＩＦ部の数を増やした場合、チャンネルＩＦ部及びディスクＩＦ部とキャッシュメモリ部との間、及びチャンネルＩＦ部及びディスクＩＦ部と共有メモリとの間のアクセスパス数も増えることになる。また、ホストコンピュータとディスクアレイ制御装置との間の接続にファイバチャネル等の高速チャネルの採用等により、ディスクアレイ制御装置に要求されるスループットはさらに増大する方向にあり、このスループットの向上の要求を満たすためには、チャンネルＩＦ部及びディスクＩＦ部とキャッシュメモリ部との間、及びチャンネルＩＦ部及びディスクＩＦ部と共有メモリとの間のアクセスパス数を増やし、内部パス性能を向上させることが必要となる。

【0011】しかし、キャッシュメモリに格納される１つのデータのデータ量は、共有メモリに格納される１つの制御情報のデータ量よりもかなり大きい。一例を挙げれば、メインフレームに接続されるディスク制御装置では、キャッシュメモリに格納される１つのデータは数Ｋバイト程度（例えば２Ｋバイト）であるのに対し、共有メモリに格納される１つの制御情報は数バイト程度（例えば４バイト）である。また、オープン系のホストコンピュータに接続されるディスク制御装置では、キャッシュメモリに格納される１つのデータは数十バイト程度（例えば６４バイト）であるのに対し、共有メモリに格納される１つの制御情報は数バイト程度（例えば４バイト）である。したがって、チャンネルＩＦ部及びディスクＩＦ部とキャッシュメモリ部との間で転送されるデータ量は、チャンネルＩＦ部及びディスクＩＦ部と共有メモリ部との間で転送されるデータ量に比べ、かなり多いので、チャンネルＩＦ部及びディスクＩＦ部とキャッシュメモリ部との間のアクセスパスのデータ幅は、チャンネルＩＦ部及びディスクＩＦ部と共有メモリ部との間のアクセスパスのデータ幅より広くとる必要がある。例えば、前者のアクセスパスは、１６ビット幅のバスで構成され、後者は、４ビット幅のバスで構成される。そのため、チャンネルＩＦ部及びディスクＩＦ部とキャッシュメモリ部との間のアクセスパスの本数を増やすと、それらのアクセスパスを接続するキャッシュメモリ部のＬＳＩのピン数が不足するという問題が生じる。また、ディスクアレイ制御装置のホストコンピュータへの応答時間を短くするためには、共有メモリ部に格納された制御情報へのアクセス時間をできるだけ短くすることも必要である。

【0012】そこで、本発明の目的は、キャッシュメモリ及び共有メモリに格納されるデータの特性及びこれら

のメモリへのアクセス特性を考慮した、スループットの高く、ディスクアレイ制御装置、及びそれを用いたサブシステムを提供することにある。

【0013】より具体的には、本発明の目的は、チャンネルＩＦ部及びディスクＩＦ部とキャッシュメモリ部との間のアクセスパスはスループットが高く、チャンネルＩＦ部及びディスクＩＦ部と共有メモリとの間のアクセスパスはスループットが高く、かつアクセス時間が短いディスクアレイ制御装置、及びそれを用いたサブシステムを提供することにある。

【0014】

【課題を解決するための手段】上記目的は、ホストコンピュータとのインターフェースを有する複数のチャンネルインターフェース部と、磁気ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、前記磁気ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリ部と、チャンネルインターフェース部及びディスクインターフェース部と前記キャッシュメモリ部との間のデータ転送に関する制御情報を格納する共有メモリ部とを有し、各チャンネルインターフェース部は、前記ホストコンピュータとのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行し、各ディスクインターフェース部は、前記磁気ディスク装置とのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行するディスクアレイ制御装置において、前記複数のチャンネルインターフェース部及び前記複数のディスクインターフェース部と前記キャッシュメモリ部との間の接続形式が、前記複数のチャンネルインターフェース部及び前記複数のディスクインターフェース部と前記共有メモリ部との間の接続形式と異なることを特徴とするディスクアレイ制御装置により達成される。

【0015】好ましくは、前記複数のチャンネルインターフェース部及び前記複数のディスクインターフェース部と前記キャッシュメモリ部との間はセクタ部を介して接続し、前記複数のチャンネルインターフェース部及び前記複数のディスクインターフェース部と前記共有メモリ部との間は、それぞれセクタ部を介せず直接接続する。

【0016】また、好ましくは、前記複数のチャンネルインターフェース部及び前記複数のディスクインターフェース部と前記キャッシュメモリ部との間はセクタ部を介して接続し、前記複数のチャンネルインターフェース部、前記複数のディスクインターフェース部、及び前記共有メモリ部は共有バスに接続する。

【0017】また、好ましくは、前記複数のチャンネルインターフェース部及び前記複数のディスクインターフェース部と前記キャッシュメモリ部との間はスイッチを用いた相互結合網によって接続し、前記複数のチャンネルインターフェース部及び前記複数のディスクインターフェ

!(7) 003-263279 (P2003-263279A)

ース部と前記共有メモリ部との間はそれぞれ直接接続する。

【0018】その他、本願が開示する課題、及びその解決方法は、発明の実施形態の欄及び図面により明らかにされる。

【0019】

【発明の実施の形態】以下、本発明の実施例を図面を用いて説明する。

【0020】[実施例1] 図1に、本発明の一実施例を示す。

【0021】ディスクアレイ制御装置1は、ホストコンピュータ50との2つのインターフェース部(チャンネルIF部)11と、磁気ディスク装置20との2つのインターフェース部(ディスクIF部)12と、2つのセクタ部13と、2つのキャッシュメモリ部14と、2つの共有メモリ部15と、アクセスパス0:135と、アクセスパス1:136と、アクセスパス2:137とを有する。

【0022】チャンネルIF部11は、ホストコンピュータ50との2つのIF(ホストIF)102と、ホストコンピュータ50に対する入出力を制御する2つのマイクロプロセッサ101と、キャッシュメモリ部14へのアクセスを制御するアクセス制御部(CMアクセス制御部)104と、共有メモリ部15へのアクセスを制御するアクセス制御部(SMアクセス制御部)105とを有し、ホストコンピュータ50とキャッシュメモリ部14間のデータ転送、及びマイクロプロセッサ101と共有メモリ部15間の制御情報の転送を実行する。マイクロプロセッサ101及びホストIF102は内部バス106によって接続され、CMアクセス制御部104は2つのホストIF102に直接接続されている。また、SMアクセス制御部105は2つのマイクロプロセッサ101に直接接続されている。

【0023】ディスクIF部12は、磁気ディスク装置20との2つのIF(ドライブIF)103と、磁気ディスク装置20に対する入出力を制御する2つのマイクロプロセッサ101と、キャッシュメモリ部14への1つのアクセス制御部(CMアクセス制御部)104と、共有メモリ部15への1つのアクセス制御部(SMアクセス制御部)105を有し、磁気ディスク装置20とキャッシュメモリ部14間のデータ転送、及びマイクロプロセッサ101と共有メモリ部15間の制御情報の転送を実行する。マイクロプロセッサ101及びドライブIF103は内部バス106によって接続され、CMアクセス制御部104は2つのドライブIF103に直接接続されている。また、SMアクセス制御部105は2つのマイクロプロセッサ101に直接接続されている。ディスクIF部はRAID機能の実行も行う。

【0024】キャッシュメモリ部14は、キャッシュメモリ(CM)コントローラ107とメモリモジュール1

09を有し、磁気ディスク装置20へ記録するデータを一時的に格納する。

【0025】共有メモリ部15は、共有メモリ(SM)コントローラ108とメモリモジュール109とを有し、ディスクアレイ制御装置1の制御情報(例えば、チャンネルIF部及びディスクIF部とキャッシュメモリ部14との間のデータ転送制御に関する情報)等を格納する。

【0026】CMアクセス制御部104には2本のアクセスパス0:135を接続し、それらを2つの異なるセクタ部13にそれぞれ接続する。セクタ部13には2本のアクセスパス1:136を接続し、それらを2つの異なるCMコントローラ107にそれぞれ接続する。したがってCMコントローラ107には、2つのセクタ部から1本ずつ、計2本のアクセスパス1:136が接続される。こうすることにより、1つのCMアクセス制御部104から1つのCMコントローラ107へのアクセスルートが2つとなる。これにより、1つのアクセスパスまたはセクタ部13に障害が発生した場合でも、もう1つのアクセスルートによりキャッシュメモリ部14へアクセスすることが可能となるため、耐障害性を向上させることができる。

【0027】SMアクセス制御部105には2本のアクセスパス2:137を接続し、そのアクセスパス2:137を2つの異なるSMコントローラにそれぞれ接続する。したがって、SMコントローラ108には、2つのチャンネルIF部11及び2つのディスクIF部12から1本ずつ、計4本のアクセスパス2:137が接続される。本実施例では、1つのSMアクセス制御部105と1つのSMコントローラ108の間には1本のアクセスパス2:137を接続したが、このアクセスパス2:137を2本に増やすことにより、1つのSMアクセス制御部105から1つのSMコントローラ108へのアクセスルートが2つとなるため、耐障害性を向上させることができる。

【0028】セクタ部13には、2つのチャンネルIF部11と、2つのディスクIF部12からそれぞれ1本ずつ、計4本のアクセスパス0:135が接続される。また、セクタ部13には、2つのキャッシュメモリ部14へのアクセスパス1:136が1本ずつ、計2本接続される。

【0029】アクセスパス0:135とアクセスパス1:136の間に上記のようなパス数の関係があるため、セクタ部13ではチャンネルIF部11及びディスクIF部12からの4本のアクセスパス0:135からの要求の内、キャッシュメモリ部14へのアクセスパス1:136の数に相当する2個だけを選択して実行する機能を持つ。

【0030】本実施例の大きな特徴は、チャンネルIF部11及びディスクIF部12とキャッシュメモリ部14

(8) 003-263279 (P2003-263279A)

との間の接続形式と、チャンネルＩＦ部１１及びディスクＩＦ部１２と共有メモリ部１５との間の接続形式とが異なる点にある。このような構成にした理由を図１及び図２を用いて説明する。図２に示したスター接続形式のディスクアレイ制御装置２では、ディスクアレイ制御装置２に搭載されるチャンネルＩＦ部１１またはディスクＩＦ部１２の数を増やした場合、チャンネルＩＦ部１１及びディスクＩＦ部１２とキャッシュメモリ部１４との間、及びチャンネルＩＦ部１１及びディスクＩＦ部１２と共有メモリ部１５との間のアクセスパス数も増えることになる。また、スター接続形式のディスクアレイ制御装置２において、スループットを向上させるためには、チャンネルＩＦ部１１及びディスクＩＦ部１２とキャッシュメモリ部１４との間、及びチャンネルＩＦ部１１及びディスクＩＦ部１２と共有メモリ部１５との間のアクセスパス数を増やし、内部パス性能を向上させることが有効である。

【００３１】しかし、キャッシュメモリ部１４に格納される１つのデータのデータ量は、共有メモリ部１５に格納される１つの制御情報のデータ量よりもかなり大きい。一例を挙げれば、メインフレームに接続されるディスク制御装置では、キャッシュメモリ部１４に格納される１つのデータは数Ｋバイト程度（例えば２Ｋバイト）であるのに対し、共有メモリ部１５に格納される１つの制御情報は数バイト程度（例えば４バイト）である。また、オープン系のホストコンピュータに接続されるディスク制御装置では、キャッシュメモリ部１４に格納される１つのデータは数十バイト程度（例えば６４バイト）であるのに対し、共有メモリ部１５に格納される１つの制御情報は数バイト程度（例えば４バイト）である。したがって、チャンネルＩＦ部１１及びディスクＩＦ部１２とキャッシュメモリ部１４との間で転送されるデータ量は、チャンネルＩＦ部１１及びディスクＩＦ部１２と共有メモリ部１５との間で転送されるデータ量に比べかなり多いので、チャンネルＩＦ部１１及びディスクＩＦ部１２とキャッシュメモリ部１４との間のアクセスパスのデータ幅は、チャンネルＩＦ部１１及びディスクＩＦ部１２と共有メモリ部１５との間のアクセスパスのデータ幅より広くとる必要がある。例えば、前者のアクセスパスは１６ビット幅のバスで構成され、後者は、８ビット幅のバスで構成される。そのため、チャンネルＩＦ部１１及びディスクＩＦ部１２とキャッシュメモリ部１４との間のアクセスパス本数を増やすと、キャッシュメモリ部１４内のキャッシュメモリコントローラ（図２ではキャッシュメモリコントローラを図示していない）のＬＳＩのピン数不足、またはキャッシュメモリ部１４を実装するパッケージにおいてコネクタのピン数不足という問題が生じる。そこで、本実施例では、図１に示すように、チャンネルＩＦ部１１及びディスクＩＦ部１２とキャッシュメモリ部１４との間をセレクト部１３を介して接続すること

により、キャッシュメモリ部１４に直接接続されるアクセスパス数を削減している。

【００３２】一方、上述したように、共有メモリ部１３へ格納する１つの制御情報のデータ長はキャッシュメモリ部１４に格納する１つのデータのデータ長に比べかなり小さいので、チャンネルＩＦ部１１及びディスクＩＦ部１２と共有メモリ部１５との間のアクセスパスのデータ幅は、チャンネルＩＦ部１１及びディスクＩＦ部１２とキャッシュメモリ部１４との間のアクセスパスのデータ幅の半分以下とすることが可能である。したがって、共有メモリ部１５へのアクセスパス数を増やしても共有メモリ部内の共有メモリメモリコントローラ（図２では共有メモリコントローラを図示していない）のＬＳＩのピン数不足等の問題が生じることは少ない。

【００３３】また、ディスクアレイ制御装置１のホストコンピュータ５０への応答時間を短くするためには、共有メモリ部１５に格納される制御情報へのアクセス時間をできるだけ短くする必要もある。しかし、図１に示したＣＭアクセス制御部１０４とＣＭコントローラ１０７間のように、ＳＭアクセス制御部１０５とＳＭコントローラ１０８との間をセレクト部を介して接続すると、セレクト部での処理のオーバーヘッドにより、共有メモリ部１５に格納される制御情報へのアクセス時間を短くすることができない。

【００３４】そこで、本実施例では、チャンネルＩＦ部１１及びディスクＩＦ部１２と共有メモリ部１５の間をセレクト部を介さず直接接続することにより、チャンネルＩＦ部１１及びディスクＩＦ部１２と共有メモリ部１５間に複数のアクセスパスを設け、スループットを向上させる一方、セレクト部での処理オーバーヘッドをなくし、チャンネルＩＦ部１１、及びディスクＩＦ部１２から共有メモリ部１５へのアクセス時間を短縮にしている。

【００３５】なお、本実施例では、耐障害性の向上という観点から、セレクト部１３、キャッシュメモリ部１４、及び共有メモリ部１５をそれぞれ二重化しているが、これらを二重化しなくても、上述の効果をえられることは言うまでもない。

【００３６】図４は、ＣＭアクセス制御部１０４内の構成を示している。ＣＭアクセス制御部１０４は、セレクト部３０２と、アドレス、コマンド、データを一時格納するパケットバッファ３０３と、セレクト部１３に繋がるアクセスパス０：１３５とのバスＩＦ３０１と、データのエラーチェック部３００と、データ転送制御部３１０を有する。セレクト部３０２の２つのポートはデータ線２１０でホストＩＦ１０２あるいはドライブＩＦ１０３に接続される。また、セレクト部３０２の他の２つのポートはバスＩＦ３０１に接続される。バスＩＦ３０１はアクセスパス０：１３５でセレクト部１３に接続される。データ転送制御部３１０は、制御線１：２１１でホストＩＦ１０２あるいはドライブＩＦ１０３に接続され、制御

!(9) 003-263279 (P2003-263279A)

線2:212でセクタ部13内のデータ転送制御部315に接続される。また、データ転送制御部310は、アービタ308によりホストIF102あるいはドライブIF103からのアクセス要求のアービトレーションを行い、セクタ302の切り替えを行う。

【0037】図6は、セクタ部13内の構成を示している。セクタ部13は、チャンネルIF部11及びディスクIF部12に繋がるアクセスパス0:135との4つのパスIF301と、CMコントローラ107に繋がるアクセスパス1:136との2つのパスIF301と、両者間を互いに接続するセクタ306と、パケットバッファ303と、データのエラーチェック部300と、CMアクセス制御部104から送出されたアドレス及びコマンドを解析するアドレス・コマンド(adr、cmd)解析部305と、データ転送制御部315を有する。データ転送制御部315は、制御線2:212でCMアクセス制御部104内のデータ転送制御部310に接続され、制御線3:213でCMコントローラ107内のデータ転送制御部315に接続される。また、データ転送制御部315は、アービタ308により、adr、cmd解析部305で解析した4本のアクセスパス0:135からのアクセス要求のアービトレーションを行い、セクタ306の切り替えを行う。パケットバッファ303は、アクセスパス0:135側のパスとアクセスパス1:136側のパスでデータ転送速度に差がある場合、速度差を吸収するために、転送するデータの一部または全部をバッファリングする。

【0038】adr、cmd解析部305は、アドレス及びコマンドを格納するバッファと、adr抽出部と、cmd抽出部を有する(図示していない)。adr、cmd解析部305では、CMアクセス制御部104に接続される4本のアクセスパス0:135それぞれに1つずつ割り当てられたバッファに、アドレス、コマンドを格納する。adr抽出部及びcmd抽出部では、アクセスするCMコントローラ107とアクセスの種類を割り出し、データ転送制御部315内のアービタ308へ送出する。

【0039】図7は、キャッシュメモリ部14内の構成を示している。キャッシュメモリ部14は、CMコントローラ107とメモリモジュール109を有する。CMコントローラ107は、セクタ部13に繋がるアクセスパス1:136との2つのパスIF301と、セクタ304と、データを一時格納するパケットバッファ303と、データのエラーチェック部300と、メモリモジュール109へのアクセスを制御するメモリ制御部307と、CMアクセス制御部104から送出されたアドレス及びコマンドを解析するadr、cmd解析部305と、データ転送制御部315を有する。データ転送制御部315は、制御線3:213でセクタ部13内のデータ転送制御部315に接続される。また、データ転

送制御部315は、アービタ308により、adr、cmd解析部305で解析した2本のアクセスパス1:136からのアクセス要求のアービトレーションを行い、セクタ304の切り替えを行う。

【0040】adr、cmd解析部305は、バッファと、adr抽出部と、cmd抽出部を有する(図示していない)。adr、cmd解析部305では、CMコントローラ107に接続される2本のアクセスパス1:136それぞれに1つずつ割り当てられたバッファに、アドレス、コマンドを格納する。adr抽出部及びcmd抽出部では、アクセスするメモリのアドレスとアクセスの種類を割り出し、メモリ制御部307へ送出する。また、2本のアクセスパス1:136からのアクセス要求をデータ転送制御部315内のアービタ308へ送出する。

【0041】次に、キャッシュメモリ部14へのアクセス時の手順について述べる。キャッシュメモリ部14へアクセスする場合、マイクロプロセッサ101は、ホストIF102あるいはドライブIF103へ、キャッシュメモリ部14へのアクセス開始を指示する。

【0042】アクセス開始の指示を受けたホストIF102あるいはドライブIF103は、制御線1:211によりCMアクセス制御部104内のデータ転送制御部310へアクセス開始を示す信号を送出する。それとともに、データ線210を通してアドレス、コマンド、データ(データの書き込み時のみ)を送出する。

【0043】CMアクセス制御部104は、データ線210を通して送られてきたアドレス、コマンド、データ(データの書き込み時のみ)をパケットバッファ303に格納する。データ転送制御部310はアービトレーションを行ってパスIF301の使用権を決定し、セクタ302を切り替える。

【0044】図9は、キャッシュメモリ部14へデータを書き込む場合の、CMアクセス制御部104からCMコントローラ107へのアクセスの流れを示している。CMアクセス制御部104内のデータ転送制御部310は、アービトレーションによってアクセスパス0:135の使用権が決定されると、制御線2:212によってセクタ部13内のデータ転送制御部315へアクセス開始を示す信号(REQ)を出す(ステップ501)。続いて、アドレス及びコマンドを送出する(ステップ502)。

【0045】セクタ部13内のデータ転送制御部315は、CMアクセス制御部104からREQ信号を受け取ると、次にアクセスパス0:135を通して送られてくるアドレス及びコマンドを受信し、adr、cmd解析部305で解析したアクセス要求に基づいてアービトレーションを行う(ステップ503)。アービトレーションの結果、アクセスパス1:136への接続権を得たら、データ転送制御部315はセクタ306を切り替

(0) 103-263279 (P 2003-263279A)

える(ステップ504)とともに、制御線2:212により、CMアクセス制御部104内のデータ転送制御部310へ、アクセスパス1:136への接続権が得られたことを示す信号(ACK)を返す(ステップ505)。次にデータ転送制御部315は、制御線3:213によってCMコントローラ107内のデータ転送制御部315へアクセス開始を示す信号(REQ)を出す(ステップ506)。続いて、アドレス及びコマンドを送出する(ステップ507)。

【0046】CMアクセス制御部104はACK信号を受けると、パケットバッファ303からデータを読み出し、セクタ302、パスIF301を介してアクセスパス0:135へ送出する。セクタ部13は、アクセスパス0:135を通して送られてきたデータを、パスIF301及びセクタ306を介してアクセスパス1:136へ送出する(ステップ509)。

【0047】CMコントローラ107内のデータ転送制御部315は、制御線3:213によってREQ信号を受け取ると、次にアクセスパス1:136を通して送られてくるアドレス及びコマンドを受信し、adr、cmd解析部305で解析したアクセス要求に基づいてアービトレーションを行い(ステップ508)、セクタ304を切り替える。アクセスパス1:136を通して送られてくるデータはパケットバッファ303に格納する。アービトレーションの結果、メモリモジュール109へのアクセス権を得たら、メモリの制御情報をメモリ制御部307へ送出し、メモリアccessのための前処理を行う(ステップ510)。次に、パケットバッファ303からデータを読み出し、セクタ304を介してメモリモジュール109へ書き込む(ステップ511)。

【0048】メモリモジュール109へのアクセスが終了すると、メモリアccessの後処理を行い、データ転送制御部315においてアクセス状況を示すステータス(STATUS)を生成する(ステップ512)。次に、ステータスをセクタ部13を介してCMアクセス制御部104へ送出する(ステップ513)。セクタ部13内のデータ転送制御部315はステータスを受け取ると、CMコントローラ107へのREQ信号をオフする(ステップ514)。CMアクセス制御部104内のデータ転送制御部310はステータスを受け取ると、セクタ部13へのREQ信号をオフする(ステップ515)。セクタ部13内のデータ転送制御部315はCMアクセス制御部104からのREQ信号のオフを確認すると、CMアクセス制御部104へのACK信号をオフする(ステップ516)。

【0049】CMアクセス制御部104内のデータ転送制御部310はステータスを受け取ると、制御線1:211により、ホストIF102あるいはドライブIF103へキャッシュメモリ部14へのアクセスの終了を報告する。

【0050】キャッシュメモリ部14からデータを読み出す場合の、CMアクセス制御部104からCMコントローラ107へのアクセスの流れは、ステップ501から508までとステップ512以降は、データの書き込みの場合と同じである。

【0051】ここでCMアクセス制御部104は、ステップ505でACK信号を受けると、データの受信待ち状態に入る。

【0052】ステップ508でメモリアccess権を得ると、CMコントローラ107はメモリモジュール109からデータを読み出し、セクタ304、パスIF301を介してアクセスパス1:136にデータを送出する。

【0053】セクタ部13は、アクセスパス1:136を通してデータを受信すると、パスIF301及びセクタ306を介してアクセスパス0:135にデータを送出する。

【0054】CMアクセス制御部104は、アクセスパス0:135を通してデータを受信すると、セクタ302、データ線210を介してホストIF102あるいはドライブIF103へデータを送出する。

【0055】図5は、SMアクセス制御部105内の構成を示している。SMアクセス制御部104は、セクタ302と、アドレス、コマンド、データを一時格納するパケットバッファ303と、SMコントローラ108に繋がるアクセスパス2:137とのパスIF301と、データのエラーチェック部300と、データ転送制御部310を有する。セクタ302の2つのポートはデータ線220でマイクロプロセッサ101に接続される。また、セクタ302の他の2つのポートはパスIF301に接続される。パスIF301はアクセスパス2:137でSMコントローラ108に接続される。データ転送制御部310は、制御線5:221でマイクロプロセッサ101に接続され、制御線6:222でSMコントローラ108内のデータ転送制御部315に接続される。また、データ転送制御部310は、アービタ308によりマイクロプロセッサ101からのアクセス要求のアービトレーションを行い、セクタ302の切り替えを行う。

【0056】図8は、共有メモリ部15内の構成を示している。共有メモリ部15は、SMコントローラ108とメモリモジュール109を有する。SMコントローラ108は、SMアクセス制御部105に繋がるアクセスパス2:137との4つのパスIF301と、セクタ309と、データを一時格納するパケットバッファ303と、データのエラーチェック部300と、メモリモジュール109へのアクセスを制御するメモリ制御部307と、SMアクセス制御部105から送出されたアドレス及びコマンドを解析するadr、cmd解析部305と、データ転送制御部315を有する。データ転送制御

(図1) 103-263279 (P2003-263279A)

部315は、制御線6:222でSMアクセス制御部105内のデータ転送制御部310に接続される。また、データ転送制御部315は、アービタ308により、adr、cmd解析部305で解析した4本のアクセスパス2:137からのアクセス要求のアービトレーションを行い、セクタ309の切り替えを行う。

【0057】adr、cmd解析部305は、バッファと、adr抽出部と、cmd抽出部を有する(図示していない)。adr、cmd解析部305では、SMコントローラ108に接続される4本のアクセスパス2:137それぞれに1つずつ割り当てられたバッファに、アドレス、コマンドを格納する。adr抽出部及びcmd抽出部では、アクセスするメモリのアドレスとアクセスの種類を割り出し、メモリ制御部307へ送出する。また、4本のアクセスパス2:137からのアクセス要求をデータ転送制御部315内のアービタ308へ送出する。

【0058】次に、共有メモリ部15へのアクセス時の手順について述べる。共有メモリ部15へアクセスする場合、マイクロプロセッサ101は、制御線5:221によりSMアクセス制御部105内のデータ転送制御部310へアクセス開始を示す信号を送出する。それとともに、データ線220を通してアドレス、コマンド、データ(データの書き込み時のみ)を送出する。

【0059】SMアクセス制御部105は、データ線220を通して送られてきたアドレス、コマンド、データ(データの書き込み時のみ)をパケットバッファ303に格納する。データ転送制御部310はアービトレーションを行ってパスIF301の使用権を決定し、セクタ302を切り替える。

【0060】図10は、共有メモリ部15へデータを書き込む場合の、SMアクセス制御部105からSMコントローラ108へのアクセスの流れを示している。SMアクセス制御部105内のデータ転送制御部310は、アービトレーションによってアクセスパス2:137の使用権が決定されると、制御線6:222によってSMコントローラ108へアクセス開始を示す信号(REQ)を出す(ステップ601)。続いて、アドレス、コマンド、及びデータを連続して送出する(ステップ602)。

【0061】SMコントローラ108内のデータ転送制御部315は、制御線6:222によってREQ信号を受け取ると、次にアクセスパス2:137を通して送られてくるアドレス、コマンド、及びデータを受信する。アドレスとコマンドは、adr、cmd解析部305で解析し、アクセス要求に基づいてアービトレーションを行い(ステップ603)、セクタ309を切り替える。データはパケットバッファ303に格納する。アービトレーションの結果、メモリモジュール109へのアクセス権を得たら、メモリの制御情報をメモリ制御部3

07へ送出し、メモリアccessのための前処理を行う(ステップ604)。次に、パケットバッファ303からデータを読み出し、セクタ309を介してメモリモジュール109へ書き込む(ステップ605)。

【0062】メモリモジュール109へのアクセスが終了すると、メモリアccessの後処理を行い、データ転送制御部315においてアクセス状況を示すステータス(STATUS)を生成する(ステップ606)。次に、ステータスをSMアクセス制御部105へ送出する(ステップ607)。SMアクセス制御部105内のデータ転送制御部310はステータスを受け取ると、SMコントローラ108へのREQ信号をオフする(ステップ608)。

【0063】SMアクセス制御部105内のデータ転送制御部310はステータスを受け取ると、制御線5:221により、マイクロプロセッサ101へ共有メモリ部15へのアクセスの終了を報告する。

【0064】共有メモリ部15からデータを読み出す場合のSMアクセス制御部105からSMコントローラ108へのアクセスの流れは、ステップ601から604までとステップ606以降は、データの書き込みの場合と同じである。

【0065】ステップ604でメモリアccessの前処理を行った後、SMコントローラ108はメモリモジュール109からデータを読み出し、セクタ309、パスIF301を介してアクセスパス2:137にデータを送出する。

【0066】SMアクセス制御部105は、アクセスパス2:137を通してデータを受信すると、セクタ302、データ線220を介してマイクロプロセッサ101へデータを送出する。

【0067】本実施例のディスクアレイ制御装置1では、ホストコンピュータ50とのチャネルを2つ有するチャネルIF部12を複数搭載しており、それらのチャネルをそれぞれ異なるホストコンピュータに接続することが可能である。そうした場合、ホストコンピュータに接続した各チャネルからの要求全てを並列に処理する必要がある。

【0068】ところでディスクアレイ制御装置1では、ホストコンピュータ50へデータを読み出す場合、磁気ディスク装置20に格納されたデータをディスクIF部12を介してキャッシュメモリ部14へ書き込み、そのデータをキャッシュメモリ部14から読み出して、チャネルIF部11を介してホストコンピュータ50へ送る。またホストコンピュータ50からディスクアレイ制御装置1へデータを書き込む場合は、ホストコンピュータ50からチャネルIF部12へ送られてきたデータをキャッシュメモリ部14へ書き込み、そのデータをキャッシュメモリ部14から読み出して、ディスクIF部12を介して磁気ディスク装置20に書き込む。さらにデ

(2) 103-263279 (P2003-263279A)

ータのパリティを生成して磁気ディスク装置20に書き込むため、ディスクIF部12とキャッシュメモリ部14間でさらに2〜3回のアクセスが行われる。

【0069】したがって、ホストコンピュータに接続した各チャンネルからの要求全てを並列に処理するためには、チャンネルIF部11及びディスクIF部12とキャッシュメモリ部14間のスループットをホストコンピュータ50とチャンネルIF部11間の最大のスループットの2倍以上にしなければならない。

【0070】本実施例では、CMアクセス制御部104とCMコントローラ107間のアクセスパスの帯域幅、及びCMコントローラ107とメモリモジュール109間の全帯域幅を、チャンネルIF部11とホストコンピュータ50間の最大の帯域幅の2倍以上に設定する。これにより、全チャンネルIF部11を並列に動作させることが可能となる。

【0071】本実施例によれば、チャンネルIF部11及びディスクIF部12とキャッシュメモリ部14間のスループットを高くすること、チャンネルIF部11及びディスクIF部12と共有メモリ部15間のスループットを高く、且つアクセス時間を短くすることの両方が可能となる。これによって、スループットが高く、且つ応答時間の短いディスクアレイ制御装置を提供できる。

【0072】ここで、図15に示すように、チャンネルIF部11及びディスクIF部12とキャッシュメモリ部14間をスイッチ(SW)16を用いた相互結合網140で接続する。この場合にも、図1に示したセクタ部13を介して接続した構成と同様に、キャッシュメモリ部14へ複数のアクセスパスを設けることができるため、スループットを高めることが可能となる。

【0073】また図16に示すように、1つのCMアクセス制御部104へ接続されるアクセスパス0:135の本数を図1の構成の倍の4本に増やしたディスクアレイ制御装置1においても、本実施例を実施する上で問題は無い。ホストIF及びドライブIFとして、今後はファイバチャネル等のスループットが100MB/s以上の高速IFが使用されることが多くなると考えられる。チャンネルIF部11及びディスクIF部12でのスループットのバランスを考えると、1つのCMアクセス制御部104に繋がる全アクセスパス0:135のスループットは、チャンネルIF部11内の全ホストIF102、またはディスクIF部12内の全ドライブIF103のスループットと同等以上にする必要がある。上記のようにファイバチャネル等の高速IFを使用する場合は、図16に示すように、1つのCMアクセス制御部104に繋がるアクセスパス0:135の本数を増やすことで、アクセスパス0:135のスループットをホストIF102あるいはドライブIF103のスループット以上にすることができる。

【0074】なお、図15及び図16のディスクアレイ

制御装置では、耐障害性の向上という観点から、セクタ部13、キャッシュメモリ部14、及び共有メモリ部15をそれぞれ二重化しているが、これらを二重化しなくても、上述の効果をえられることは言うまでもない。

【0075】〔実施例2〕図1において、2つのキャッシュメモリ部14間で、メモリ領域の全部または、一部を二重化し、キャッシュメモリ部14へのデータの書き込み時に二重化した2つの領域に同じデータを書き込むことにより、データの信頼性を上げることが可能となる。

【0076】二重化した2つのキャッシュメモリ部14へデータを書き込む場合の手順は以下のようになる。

【0077】アクセス開始の指示を受けたホストIF102あるいはドライブIF103は、制御線1:211によりCMアクセス制御部104内のデータ転送制御部310へアクセス開始を示す信号を送出する。それとともに、データ線210を通して2つのアドレス、コマンドと、1つのデータを送出する。

【0078】CMアクセス制御部104は、データ線210を通して送られてきた2つのアドレス、コマンドと、1つのデータをパケットバッファ303に格納する。データ転送制御部310はアービトレーションを行ってバスIF301の使用権を決定し、セクタ302を切り替える。

【0079】図11は、二重化した2つのキャッシュメモリ部14へデータを書き込む場合の、CMアクセス制御部104から2つのCMコントローラa、b:107へのアクセスの流れを示している。CMアクセス制御部104内のデータ転送制御部310は、アービトレーションによってアクセスパス0:135の使用権が決定されると、制御線2:212によってセクタ部13内のデータ転送制御部315へアクセス開始を示す信号(REQ)を出す(ステップ701)。続いて、アドレス及びコマンドを2つ連続して送出的(ステップ702)。

【0080】セクタ部13内のデータ転送制御部315は、CMアクセス制御部104からREQ信号を受け取ると、次にアクセスパス0:135を通して送られてくるアドレス及びコマンドを受信し、adr、cmd解析部305で解析したアクセス要求に基づいてアービトレーションを行う(ステップ703)。アービトレーションの結果、CMコントローラa、b:107への2本のアクセスパス1:136への接続権の両方を得たら、データ転送制御部315はセクタ306を切り替える(ステップ704)とともに、制御線2:212により、CMアクセス制御部104内のデータ転送制御部310へ、CMコントローラa、b:107への接続権が得られたことを示す信号(ACK)を返す(ステップ705)。次にデータ転送制御部315は、制御線3:213によって2つのCMコントローラ107内のデータ

(図3) 103-263279 (P2003-263279A)

転送制御部315へアクセス開始を示す信号(REQ)を出す(ステップ706a、706b)。続いて、CMコントローラa、b:107へアドレス及びコマンドを1つずつ送出する(ステップ707a、707b)。

【0081】CMアクセス制御部104はACK信号を受けると、パケットバッファ303からデータを読み出し、セクタ302、バスIF301を介してアクセスバス0:135へ送出する(ステップ709)。セクタ部13は、アクセスバス0:135を通して送られてきた1つのデータを、バスIF301及びセクタ306を介して2つのアクセスバス1:136の両方へ送出する(ステップ709a、709b)。

【0082】CMコントローラa、b:107内のデータ転送制御部315は、制御線3:213によってREQ信号を受け取ると、次にアクセスバス1:136を通して送られてくるアドレス及びコマンドを受信し、addr、cmd解析部305で解析したアクセス要求に基づいてアービトレーションを行い(ステップ708a、708b)、セクタ304を切り替える。アクセスバス1:136を通して送られてくるデータはパケットバッファ303に格納する。アービトレーションの結果、メモリモジュール109へのアクセス権を得たら、メモリの制御情報をメモリ制御部307へ送出し、メモリアクセスのための前処理を行う(ステップ710a、710b)。次に、パケットバッファ303からデータを読み出し、セクタ304を介してメモリモジュール109へ書き込む(ステップ711a、711b)。

【0083】メモリモジュール109へのアクセスが終了すると、メモリアクセスの後処理を行い、データ転送制御部315においてアクセス状況を示すステータス(STATUS)を生成する(ステップ712a、712b)。次に、ステータスをセクタ部13を介してCMアクセス制御部104へ送出する(ステップ713a、713b)。セクタ部13内のデータ転送制御部315はステータスを受け取ると、CMコントローラa、b:107へのREQ信号をそれぞれオフする(ステップ714a、714b)。また、セクタ部13はCMコントローラa、b:107の両方からステータスを受け取ったら、それらを続けてCMアクセス制御部へ送出する(ステップ713)。CMアクセス制御部104内のデータ転送制御部310は2つのステータスを受け取ると、セクタ部13へのREQ信号をオフする(ステップ715)。セクタ部13内のデータ転送制御部315はCMアクセス制御部104からのREQ信号のオフを確認すると、CMアクセス制御部104へのACK信号をオフする(ステップ716)。

【0084】CMアクセス制御部104内のデータ転送制御部310はステータスを受け取ると、制御線1:211により、ホストIF102あるいはドライブIF103へキャッシュメモリ部14へのアクセスの終了を報

告する。

【0085】上記のように二重化したキャッシュメモリ部14にデータを二重に書き込む場合、二重に書き込むデータの一方のデータの書き込み処理を行っている間に、もう一方のデータが他のアクセスバスからの書き込み要求によって書き換えられるのを防ぐ必要がある。本実施例では、共有メモリ部15にキャッシュメモリ部14のディレクトリを格納し、キャッシュメモリ部14にアクセスする前に必ず、共有メモリ部15に格納したディレクトリにアクセス中を示すビットを立てる。これにより、キャッシュメモリ部14内の同じアドレスには同時に1つのアクセス要求しか発行されないため、二重に書き込むデータの一方のデータの書き込み処理を行っている間に、もう一方のデータが他のアクセスバスからの書き込み要求によって書き換えられるのを防ぐことができる。

【0086】[実施例3] ディスクアレイ制御装置1では、キャッシュメモリ部14を複数設けた場合、あるキャッシュメモリ部14から別のキャッシュメモリ部14へデータをコピーする機能が要求される。この機能は、以下に述べる手順で実現できる。

【0087】アクセス開始の指示を受けたホストIF102あるいはドライブIF103は、制御線1:211によりCMアクセス制御部104内のデータ転送制御部310へアクセス開始を示す信号を送出する。それとともに、データ線210を通して2つのアドレス、コマンドを送出する。2つのうちの1つのアドレス及びコマンドは、コピー元のアドレスとリードコマンドで、もう1つのアドレス及びコマンドはコピー先のアドレスとライトコマンドである。ここでは、CMコントローラa:107をコピー元、CMコントローラbをコピー先として説明する。

【0088】CMアクセス制御部104は、データ線210を通して送られてきた2つのアドレス、コマンドをパケットバッファ303に格納する。データ転送制御部310はアービトレーションを行ってバスIF301の使用権を決定し、セクタ302を切り替える。

【0089】図12は、2つのキャッシュメモリ部間でデータをコピーする場合の、CMアクセス制御部104からCMコントローラa、b:107へのアクセスの流れを示している。CMアクセス制御部104内のデータ転送制御部310は、アービトレーションによってアクセスバス0:135の使用権が決定されると、制御線2:212によってセクタ部13内のデータ転送制御部315へアクセス開始を示す信号(REQ)を出す(ステップ801)。続いて、アドレス及びコマンドを2つ連続して送出する(ステップ802)。

【0090】セクタ部13内のデータ転送制御部315は、CMアクセス制御部104からREQ信号を受け取ると、次にアクセスバス0:135を通して送られて

(4) 103-263279 (P 2003-263279A)

くるアドレス及びコマンドを受信し、adr、cmd解析部305で解析したアクセス要求に基づいてアービトレーションを行う(ステップ803)。アービトレーションの結果、CMコントローラa、b:107への2つのアクセスパス1:136への接続権の両方を得たら、データ転送制御部315はセクタ306を切り替える(ステップ804)とともに、制御線2:212により、CMアクセス制御部104内のデータ転送制御部310へ、CMコントローラa、b:107両方への接続権が得られたことを示す信号(ACK)を返す(ステップ805)。次にデータ転送制御部315は、制御線3:213によってCMコントローラa、b:107内のデータ転送制御部315へアクセス開始を示す信号(REQ)を出す(ステップ806a、806b)。続いて、CMコントローラa、b:107へそれぞれのアドレス及びコマンドを送出する(ステップ807a、807b)。

【0091】CMアクセス制御部104はACK信号を受けると、アクセスの終了を知らせるステータスの受信待ち状態に入る。

【0092】コピー元のCMコントローラa:107内のデータ転送制御部315は、制御線3:213によってREQ信号を受け取ると、次にアクセスパス1:136を通して送られてくるアドレス及びコマンドを受信し、adr、cmd解析部305で解析したアクセス要求に基づいてアービトレーションを行い(ステップ808)、セクタ304を切り替える。アービトレーションの結果、メモリモジュール109へのアクセス権を得たら、メモリの制御情報をメモリ制御部307へ送出し、メモリアクセスのための前処理を行う(ステップ809)。次に、メモリモジュール109からデータを読み出し(ステップ810)、セクタ304を介してアクセスパス1:136へ送出する(ステップ811a)。

【0093】セクタ部13は、アクセスパス1:136を通してCMコントローラa:107から送られてきたデータをCMコントローラbに繋がるアクセスパス1:136へ送出する。(ステップ811b)。

【0094】CMコントローラb:107内のデータ転送制御部315は、制御線3:213によってREQ信号を受け取ると、次にアクセスパス1:136を通して送られてくるアドレス及びコマンドを受信し、adr、cmd解析部305で解析する。その後、データを受信し始めるのを待って、メモリアクセスのアービトレーションに参加する(ステップ812)。アクセスパス1:136を通して送られてくるデータはパケットバッファ303に格納する。アービトレーションの結果、メモリモジュール109へのアクセス権を得たら、メモリの制御情報をメモリ制御部307へ送出し、メモリアクセスのための前処理を行う(ステップ813)。次に、パケ

ットバッファ303からデータを読み出し、セクタ304を介してメモリモジュール109へ書き込む(ステップ814)。

【0095】CMコントローラa、b:107は、それぞれCMメモリモジュール109へのアクセスが終了すると、メモリアクセスの後処理を行い、データ転送制御部315においてアクセス状況を示すステータス(STATUS)を生成する(ステップ815、818)。次に、ステータスをセクタ部13へ送出する(ステップ816、819)。

【0096】セクタ部13内のデータ転送制御部315はステータスを受け取ると、CMコントローラa、b:107へのREQ信号をそれぞれオフする(ステップ817、821)。また、セクタ部13は2つのCMコントローラa、b:107の両方からステータスを受け取ったら、それらを続けてCMアクセス制御部へ送出する(ステップ820)。CMアクセス制御部104内のデータ転送制御部310は2つのステータスを受け取ると、セクタ部13へのREQ信号をオフする(ステップ822)。セクタ部13内のデータ転送制御部315はCMアクセス制御部104からのREQ信号のオフを確認すると、CMアクセス制御部104へのACK信号をオフする(ステップ823)。

【0097】CMアクセス制御部104内のデータ転送制御部310はステータスを受け取ると、制御線1:211により、ホストIF102あるいはドライブIF103へキャッシュメモリ部14へのアクセスの終了を報告する。

【0098】ディスクアレイ制御装置1ではまた、1つのキャッシュメモリ部14内のあるアドレスから別のアドレスへデータをコピーする機能も要求される。

【0099】この機能は、図9で示したデータの書き込み時の手順において、ステップ511のメモリモジュール109へのライトアクセスの代わりに、メモリモジュール109からデータを読み出してCMコントローラ107内のパケットバッファ303に格納し、続けてそのデータをメモリモジュール109へ書き込むという処理を行うことによって実現できる。

【0100】[実施例4]図1において、2つの共有メモリ部15間で、メモリ領域の全部または、一部を二重化し、共有メモリ部15へのデータの書き込み時に二重化した2つの領域に同じデータを書き込むことにより、データの信頼性を上げることが可能となる。

【0101】二重化した2つの共有メモリ部15へデータを書き込む場合の手順は、以下になる。

【0102】マイクロプロセッサ101は、制御線5:221によりSMアクセス制御部105内のデータ転送制御部310へアクセス開始を示す信号を送出する。それとともに、データ線220を通して2つのアドレス、コマンドと、1つのデータを送出する。

(図 5) 103-263279 (P 2003-263279A)

【0103】SMアクセス制御部105は、データ線20を通して送られてきた2つのアドレス、コマンドと、1つのデータをバケットバッファ303に格納する。データ転送制御部310はアービトレーションを行ってバスIF301の使用権を決定し、セクタ302を切り替える。

【0104】図13は、二重化した2つの共有メモリ部15へデータを書き込む場合の、SMアクセス制御部105から2つのSMコントローラ108へのアクセスの流れを示している。2つの共有メモリ部を二重化する場合、一方をマスタ、もう一方をスレーブに設定する。SMアクセス制御部105内のデータ転送制御部310は、アービトレーションによって、まずマスタ側のSMコントローラ108へのアクセスパス2:137の使用権を決定し、制御線6:222によってマスタ側のSMコントローラ108内のデータ転送制御部315へアクセス開始を示す信号(REQ)を出す(ステップ901)。続いて、アドレス、コマンド、及びデータを連続して送出する(ステップ902)。

【0105】マスタ側のSMコントローラ108内のデータ転送制御部315は、制御線6:222によってREQ信号を受け取ると、次にアクセスパス2:137を通して送られてくるアドレス、コマンド、及びデータを受信し、adr、cmd解析部305で解析したアクセス要求に基づいてアービトレーションを行い(ステップ903)、セクタ309を切り替える。データはバケットバッファ303に格納する。アービトレーションの結果、メモリモジュール109へのアクセス権を得たら、メモリの制御情報をメモリ制御部307へ送出し、メモリアクセスのための前処理を行う(ステップ904)。次に、バケットバッファ303からデータを読み出し、セクタ309を介してメモリモジュール109へ書き込む(ステップ905)。

【0106】メモリモジュール109へのアクセスが終了すると、メモリアクセスの後処理を行い、データ転送制御部315においてアクセス状況を示すステータス(STATUS)を生成する(ステップ906)。次に、ステータスをSMアクセス制御部105へ送出する(ステップ907)。

【0107】SMアクセス制御部105はステータスを受け取ったら、マスタ側のSMコントローラ108へのアクセスパス2:137の使用権を開放せずに、アービトレーションによって、スレーブ側のSMコントローラ108へのアクセスパス2:137の使用権を決定する。その後のスレーブ側のSMコントローラ108へのアクセス手順(ステップ908~914)は、マスタ側のSMコントローラ108へのアクセス手順(ステップ901~907)と同様である。

【0108】SMアクセス制御部105内のデータ転送制御部310はスレーブ側のSMコントローラ108か

らステータスを受け取ると、マスタ側とスレーブ側両方のSMコントローラ108へのREQ信号をオフする(ステップ715a、715b)。

【0109】二重化した共有メモリ部15にデータを二重に書き込む場合、上記のように共有メモリ部15をマスタとスレーブに分け、マスタ側、スレーブ側の順でデータを書き込み、スレーブ側へのデータへの書き込みが終わるまで両方のアクセスパス2:137を開放しない。これによりデータを書き込む順番が保証され、二重に書き込むデータの一方のデータの書き込み処理を行っている間に、もう一方のデータが他のアクセスパスからの書き込み要求によって書き換えられるのを防ぐことができる。

【0110】SMアクセス制御部105内のデータ転送制御部310はステータスを受け取ると、制御線5:221により、マイクロプロセッサ101へ共有メモリ部15へのアクセスの終了を報告する。

【0111】[実施例5]図14に、実施例1のディスクアレイ制御装置1のチャンネルIF部11、ディスクIF部12、セクタ部13、キャッシュメモリ部14、共有メモリ部15を実装するときの構成を示す。

【0112】チャンネルIF部11、ディスクIF部12、セクタ部13、キャッシュメモリ部14、共有メモリ部15は、それぞれ独立したパッケージ(PK)、すなわち、チャンネルIFPK1、ディスクIFPK2、セクタPK3、メモリPK4にそれぞれ実装する。異なるキャッシュメモリ部14は、異なるパッケージ上に実装する。また、異なる共有メモリ部15も、異なるパッケージ上に実装する。1つのキャッシュメモリ部14と1つの共有メモリ部15は同一のパッケージ上に実装しても問題ない。図14では、1つのキャッシュメモリ部14と1つの共有メモリ部15を同一のメモリPK4上に実装した例を示している。

【0113】チャンネルIFPK1、ディスクIFPK2、セクタPK3、メモリPK4は、ブラッタ5上に実装し、それを筐体に搭載する。

【0114】ここで、各PKをブラッタ5に実装するときには、セクタPK3を中心付近より外側、好ましくはブラッタの両端に配置することが重要となる。このような配置により、各PK間を結ぶ線をブラッタ5上に配線する際、ブラッタ5全体にわたって配線の密度を均一にでき、ブラッタ上の配線を容易になる。

【0115】本実施例では、セクタ部13をセクタPK3に実装して、ブラッタ5の両端に配置するとした。しかし、セクタ部13をパッケージに実装せず、ブラッタの両端に直接実装しても問題ない。

【0116】[実施例6]図1に示す実施例1のディスクアレイ制御装置1において、SMアクセス制御部105とSMコントローラ108間をアクセスパス2:137で接続する代わりに、図7に示すように、2本の共有

(図6) 103-263279 (P2003-263279A)

バス130を介して接続する。SMアクセス制御部105からは、2本の共有バス130それぞれに1本ずつ接続バスを設ける。また、SMコントローラ108からも、2本の共有バス130それぞれに1本ずつ接続バスを設ける。SMアクセス制御部105及びSMコントローラ108は、それぞれ2つのアービタを有している。2つのアービタはそれぞれ2本の共有バスのアービトレーション用のアービタである。SMアクセス制御部105からSMコントローラ108へのアクセスの際は、複数のアービタの内の1つがマスタとなり、共有バス130の使用権のアービトレーションを行う。そして、使用権を得たSMアクセス制御部105がSMコントローラ108にアクセスを行う。また、アービタをSMアクセス制御部105及びSMコントローラ108内に設ける代わりに、独立した回路として共有バス130に直接接続しても問題ない。

【0117】共有バス接続では、バスのデータ幅を広げることによりデータ転送速度を上げることが可能であり、共有メモリ部15へのアクセス時間を短縮可能である。

【0118】上述したように、図1におけるアクセスバス2:137のデータ幅は、アクセスバス0:135のデータ幅よりも2倍以上小さくすることができるので、図1のようにチャンネルIF部11及びディスクIF部12と共有メモリ部15との間をスター接続(1対1接続)しても、共有メモリ部を実装するLSIのピンネックの問題が生じるケースは少ない。とはいえ、アクセスバス2:137の本数が増え過ぎてアクセスバス2:137を実装できないという問題が生じる可能性もある。そうした場合、本実施例の共有バス接続が有効となる。

【0119】なお、本実施例では、耐障害性の向上という観点から、セクタ部13、キャッシュメモリ部14、及び共有メモリ部15をそれぞれ二重化しているが、これらを二重化しなくても、上述の効果をえられることは言うまでもない。

【0120】

【発明の効果】本発明によれば、チャンネルIF部、ディスクIF部-キャッシュメモリ間のアクセスパスについてはスループットを高くできる。また、チャンネルIF部、ディスクIF部-共有メモリ間のアクセスパスについてはスループットを高くでき、かつアクセス時間を短くできる。これによって、スループットが高く、かつ応答時間の短いディスクアレイ制御装置を提供することが可能となる。

【図面の簡単な説明】

【図1】本発明によるディスクアレイ制御装置の構成を示す図。

【図2】従来のディスクアレイ制御装置の構成を示す図。

【図3】従来のディスクアレイ制御装置の他の構成を示す図。

【図4】本発明によるディスクアレイ制御装置内のCMアクセス制御部の構成を示す図。

【図5】本発明によるディスクアレイ制御装置内のSMアクセス制御部の構成を示す図。

【図6】本発明によるディスクアレイ制御装置内のセクタ部の構成を示す図。

【図7】本発明によるディスクアレイ制御装置内のキャッシュメモリ部の構成を示す図。

【図8】本発明によるディスクアレイ制御装置内の共有メモリ部の構成を示す図。

【図9】キャッシュメモリ部へのデータの書き込み時の手順を示す図。

【図10】共有メモリ部へのデータの書き込み時の手順を示す図。

【図11】二重化した2つのキャッシュメモリ部ヘデータを二重に書き込む時の手順を示す図。

【図12】1つのキャッシュメモリ部から別のキャッシュメモリ部ヘデータをコピーする時の手順を示す図。

【図13】二重化した2つの共有メモリ部ヘデータを二重に書き込む時の手順を示す図。

【図14】本発明によるディスクアレイ制御装置内の実装の構成を示す図。

【図15】本発明によるディスクアレイ制御装置内の他の構成を示す図。

【図16】本発明によるディスクアレイ制御装置内の他の構成を示す図。

【図17】本発明によるディスクアレイ制御装置の構成を示す図。

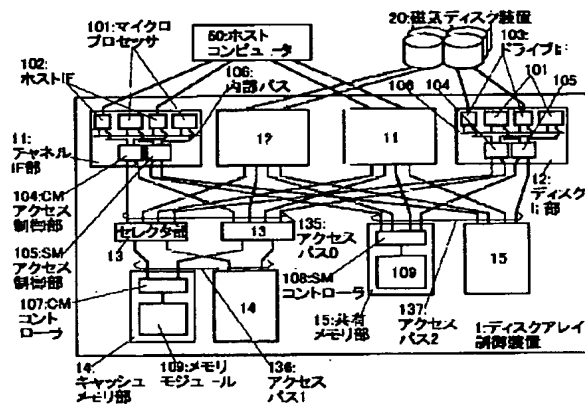
【符号の説明】

1…ディスクアレイ制御装置、11…チャンネルIF部、12…ディスクIF部、13…セクタ部、14…キャッシュメモリ部、15…共有メモリ部、20…磁気ディスク装置、50…ホストコンピュータ、101…マイクロプロセッサ、102…ホストIF、103…ドライバIF、104…CMアクセス制御部、105…SMアクセス制御部、106…内部バス、107…CMコントローラ、108…SMコントローラ、109…メモリモジュール、135…アクセスバス0、136…アクセスバス1、137…アクセスバス2。

(7) 03-263279 (P 2003-263279A)

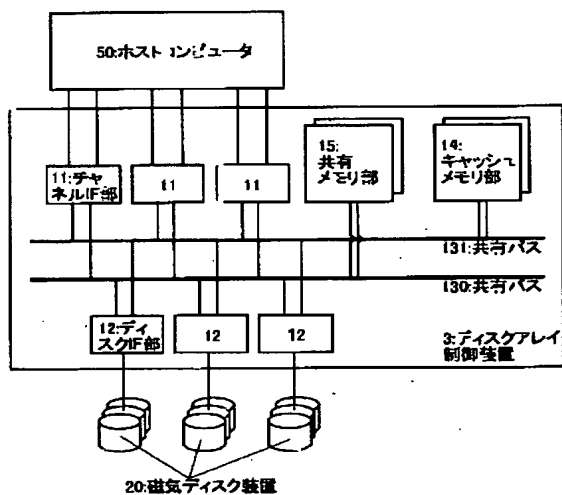
【図1】

図1



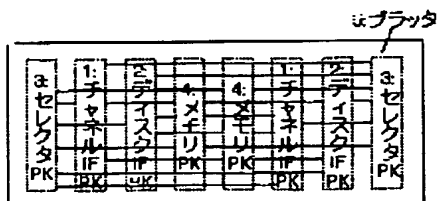
【図3】

図3



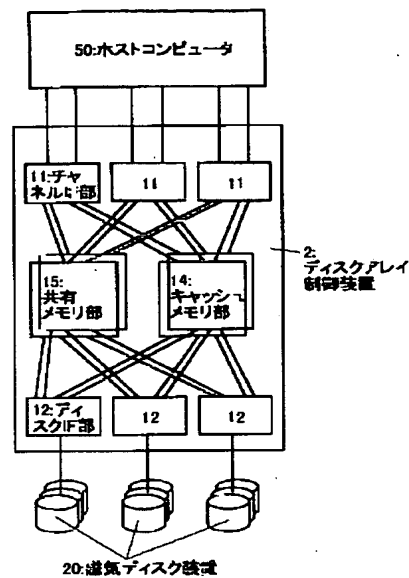
【図14】

図14



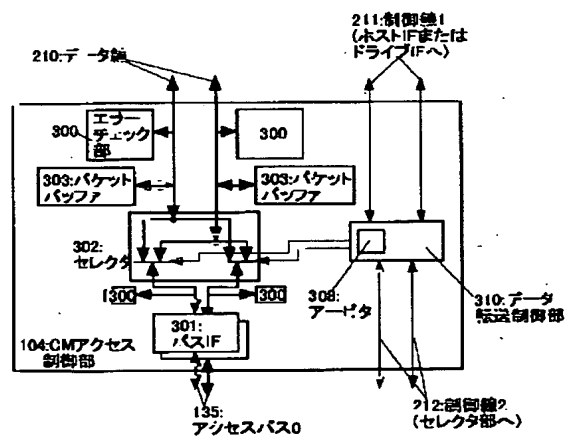
【図2】

図2



【図4】

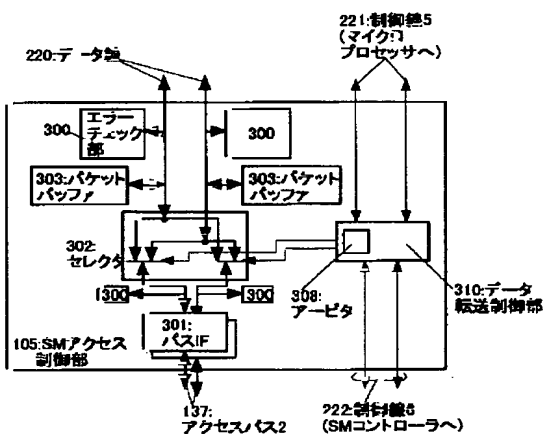
図4



(直8) 103-263279 (P2003-263279A)

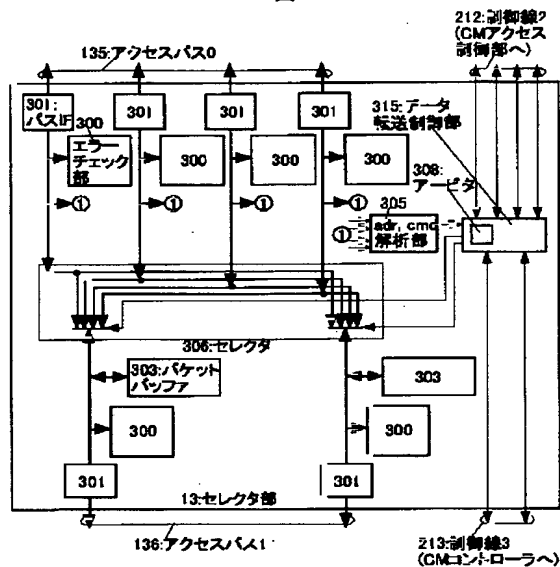
【図5】

図5



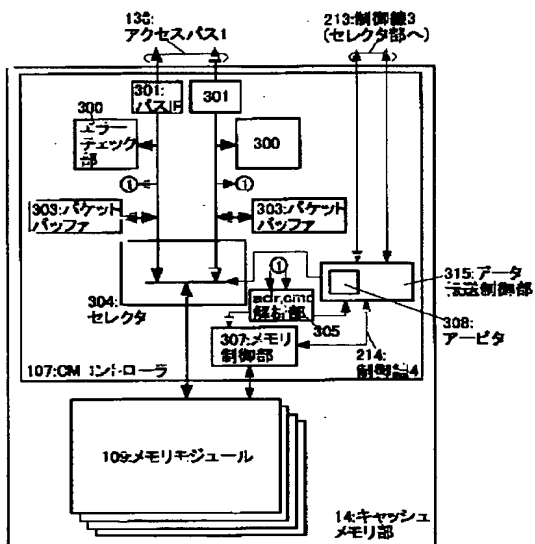
【図6】

図6



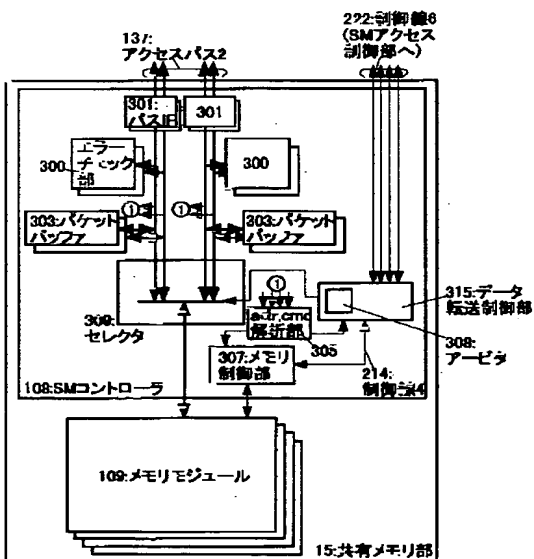
【図7】

図7



【図8】

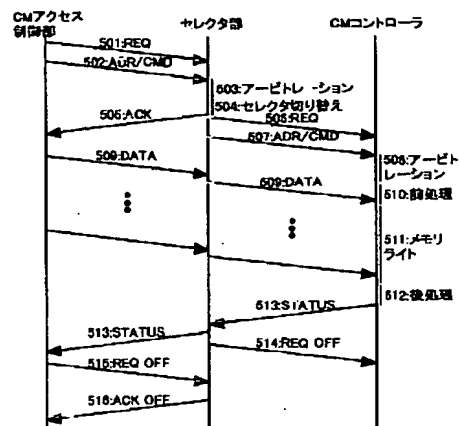
図8



(株) 103-263279 (P2003-263279A)

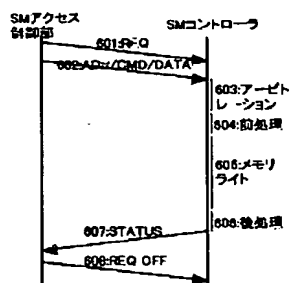
【図9】

図9



【図10】

図10

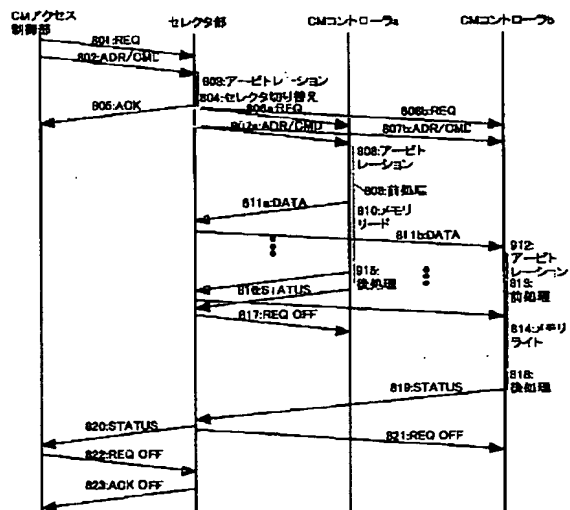
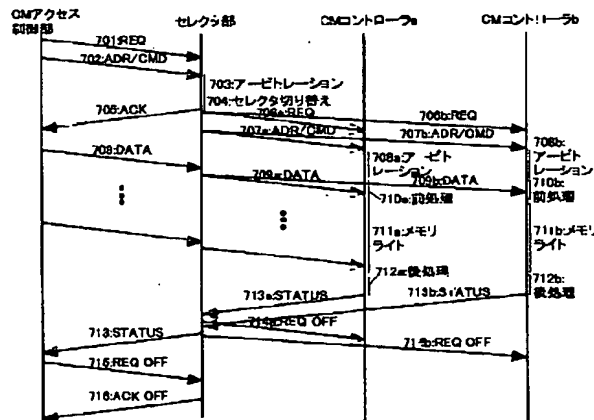


【図12】

図12

【図11】

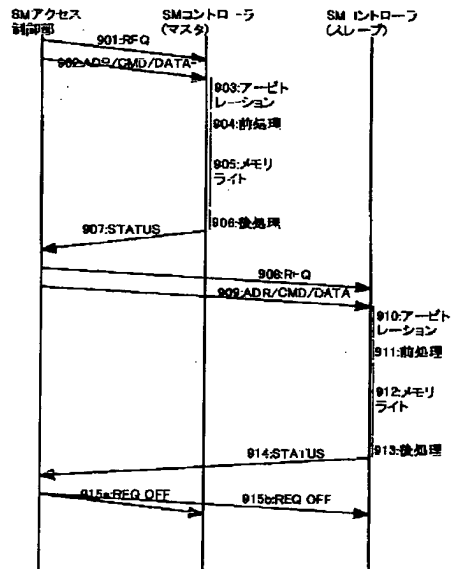
図11



(20) 103-263279 (P2003-263279A)

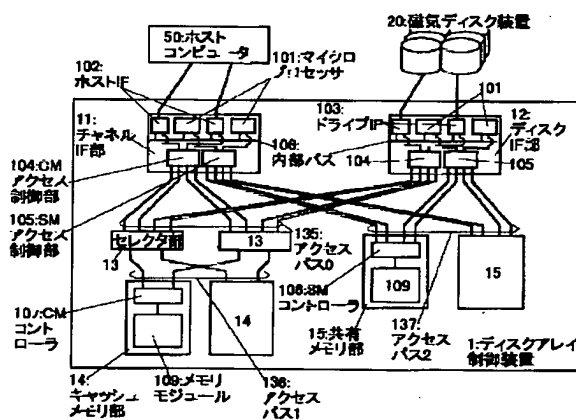
【図13】

図13



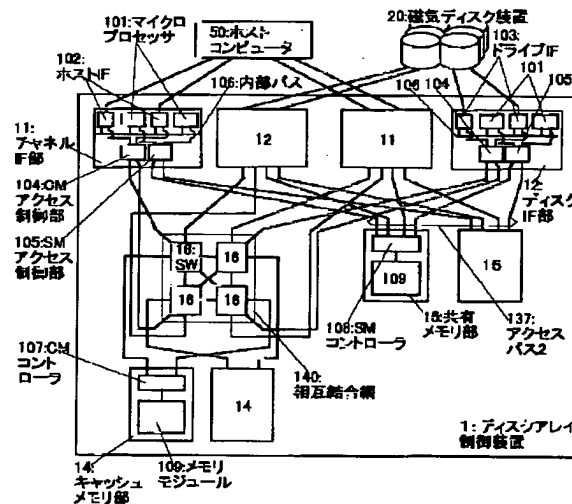
【図16】

図16



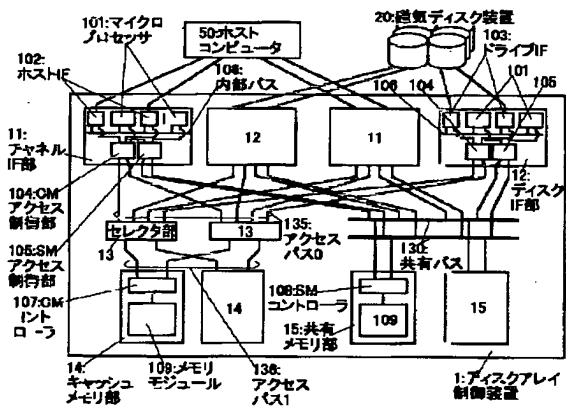
【図15】

図15



【図17】

図17



フロントページの続き

(72)発明者 藤林 昭

東京都国分寺市東恋ヶ窪一丁目280番地
株式会社日立製作所中央研究所内

(72)発明者 金井 宏樹

東京都国分寺市東恋ヶ窪一丁目280番地
株式会社日立製作所中央研究所内

(72)発明者 箕輪 信幸

神奈川県小田原市国府津2880番地 株式会
社日立製作所ストレージシステム事業部内

F ターム(参考) 5B005 JJ01 JJ12 MM12 NN12 WW12

5B065 BA01 CA11 CA30 CE14 CH01